Host plant identification in the generalist xylem feeder Philaenus spumarius through gut content analysis

Host plant identification in the generalist xylem feeder *Philaenus spumarius* through gut content analysis

Gonella Elena[1]*, Luca Picciau[1], Liam Pippinato[1], Beniamino Cavagna[2], Alberto Alma[1]

[1]Dipartimento di Scienze Agrarie, Forestali e Alimentari (DISAFA), Università degli Studi di Torino, Grugliasco, Italy

[2]Plant Protection Service, Regione Lombardia, Milano, Italy

Corresponding author: elena.gonella@unito.it, Phone +39 011 6708700, Fax +39 011 6708535

Running Title

Gut content analysis of *Philaenus spumarius*

Keywords

Meadow spittlebug, vector, polyphagy, adults, chloroplast DNA, *Xylella fastidiosa*

Abstract

The meadow spittlebug *Philaenus spumarius* (Hemiptera: Aphrophoridae) is the main vector of the phytopathogenic bacterium *Xylella fastidiosa* in Europe, where the ST53 strain induces the olive quick decline syndrome causing the most severe economic damage in southern Italy. The wide range of plant species infected by *X. fastidiosa*, and the wide host range of *P. spumarius* suggest that a huge number of wild and cultivated plants may become infected by the pathogen following unintentional introduction events. Therefore, it is necessary to detail the host plant preference of the vector, in order to include preferred in-field plants in pathogen-targeted diagnostic efforts. This will allow the identification of main sources for *X. fastidiosa* acquisition by *P. spumarius*; such plant species represent an important target for rational disease management. Here we investigated the host plants of *P. spumarius* in north-western Italy, a region where *X. fastidiosa* is still not present but is regarded as a primary threat. We designed a new molecular diagnostic tool targeting chloroplast DNA, to characterize the gut content of single *P. spumarius* adults. The newly set up nested PCR/sequencing-based identification protocol was proven to be useful for retrieving sequences from the two last different host-plant used by *P. spumarius*, even if a limited persistence of intact chloroplast DNA was reported in the spittlebug gut. We propose this protocol as a new tool for supporting research on xylem feeders biology that could be particularly useful for highly polyphagous species like *P. spumarius*. Furthermore, the method could assist monitoring of *X. fastidiosa* invasion, by contributing to the study of vector ecology and pathogen epidemics.

Introduction

The plant pathogen *Xylella fastidiosa* is a Gram negative bacterium belonging to Gammaproteobacteria (Chatterjee et al., 2008). It resides in xylem vessels of infected plants, inducing serious symptoms related to the occlusion of xylem vessels due to bacterial colonization and plant response (Sicard et al., 2018). Three distinct subspecies have been recognized based on genomic identity, namely *fastidiosa*, *multiplex* and *pauca* (Marcelletti & Scortichini, 2016). The subspecies *fastidiosa* causes one of the most troubling diseases, Pierce's Disease (PD) of grapevine, as well as leaf scorch diseases to coffee and oleander, whereas the subspecies *multiplex* is related to almond leaf scorch (Baldi & La Porta, 2017). *X. fastidiosa* ssp. *pauca* strains are the agents of citrus variegated chlorosis, a severe disease widespread in South American citrus producing areas (Cordeiro et al., 2014), and of olive quick decline syndrome, which rapidly devastated olive production in Apulia region of Italy, demonstrating the huge damaging potential of this pathogen in Europe (Strona et al., 2017). However, *X. fastidiosa* is widely generalist: its host range encompasses more than 560 plant species, with a rapid increase of reports on new infected plants following its spread in Europe from 2013 (EFSA, 2018). Its capability to infect a large number of plants is furtherly exacerbated by its transmission. Indeed, *X. fastidiosa* is transmitted by xylem feeding hemipterans, within the infraorder Cicadomorpha, belonging to the families Cicadellidae (subfamily Cicadellinae), Cercopidae, Cicadidae and Aphrophoridae (Redak et al., 2004). Low specificity has been reported for the interaction between *X. fastidiosa* strains and their vector species, and potentially all xylem feeders can be vectors (Almeida et al., 2005). Moreover, many vectors are highly polyphagous and widely distributed, being capable to easily move among habitats (Redak et al. 2004; Cornara et al., 2018a; Krugner et al., 2019), further increasing the dissemination potential of *X. fastidiosa*.

Since the first detection of *X. fastidiosa* in olive trees in Italy in 2013, relevant resources have been employed to study the spread of this pathogen and the consequent disease outbreaks in Europe and

63 in the Mediterranean basin, where *X. fastidiosa* is considered one of the most dangerous agricultural
64 threats. Hence, the intensification of studies regarding *X. fastidosa*-related epidemics is required for
65 designing effective control strategies (Sicard et al., 2018). The main vector of *X. fastidiosa* in Italy
66 is *Philaenus spumarius* L. (Hemiptera: Aphrophoridae), which is a cosmopolitan and highly
67 polyphagous species, having the potential to widely expand the pathogen distribution in Europe
68 (Cornara et al., 2018a). Since *P. spumarius* can feed on a broad range of monocotyledons and
69 dicotyledons (Dongiovanni et al., 2019), and the plant composition at the landscape scale is
70 predicted to affect its spatial distribution (Santoiemma et al., 2019), understanding the feeding
71 preference of vectors is a crucial issue for the management of *X. fastidiosa*.

72 The molecular analysis of gut content has been widely applied to delineate cryptic trophic
73 behaviour of insects (Pompanon et al., 2012); PCR-based techniques have been commonly used to
74 assess predator-prey and parasitoid-host relations (Sheppard & Harwood, 2005; Gariepy et al.,
75 2007). In herbivorous insects, the molecular analysis of gut content has been proposed as a method
76 for elucidating multiple plant use by single individuals (Hereward & Walter, 2012). Most of the
77 work has been focused on chewing species (Matheson et al., 2008; Jurado-Rivera et al., 2009;
78 Pumariño et al., 2011; Avanesyan, 2014; De la Cadena et al., 2017), as a large amount of
79 chloroplast DNA – the target for molecular analyses – can be retrieved from their gut. Conversely,
80 analysing the gut content of sap-feeding insects may be considerably challenging, considering the
81 low DNA load of plant sap. However, recently Rodney Cooper et al. (2016) successfully detected
82 plant DNA from the phloem feeder *Bactericera cockerelli* (Sulc) (Hemiptera: Triozidae). It has
83 been suggested that plant DNA may be ingested by sap feeders following stylet contamination
84 during penetration into the parenchyma cells before reaching vascular tissues (Pearson et al., 2014;
85 Rodney Cooper et al., 2016). Hence, not only phloem feeders but also xylem feeders may be
86 exposed to plant DNA during probing of plant tissue, since during this phase the stylets cross
87 periderm and parenchyma cells (Miranda et al., 2009; Cornara et al., 2018b). However, at present

88 the gut content of xylem feeding insects has never been molecularly investigated to search for plant
89 DNA.

90 In this study, we tracked the presence of plant-related DNA in the gut of *P. spumarius*, by
91 developing a specific protocol for xylem feeders, to identify host plants used by adult individuals of
92 *P. spumarius*. We evaluated the persistence of target chloroplast DNA in the digestive tract of the
93 spittlebug over a period of three days in the absence of plant substrate or after a plant change. The
94 aim of this work was to assess the time range allowing molecular detection of plant-related DNA in
95 the gut content of a polyphagous xylem feeder. Such a molecular tool could provide a support to
96 classical bioassays in improving the knowledge on its plant host range, to investigate its seasonal
97 movement based on food sources availability. Moreover, tracking the diet of single spittlebug
98 individuals could be applied to all the other vector species recognized in Europe – which are close
99 relatives of *P. spumarius* (Cavalieri et al., 2019) – assisting *X. fastidiosa* monitoring efforts in
100 vectors. Indeed, the identification of the last host plants used by infected specimens will allow
101 recognizing possible infection reservoirs, since no latency is needed for *X. fastidosa* transmission by
102 *P. spumarius* (Cornara et al., 2016).

103

104 ## Materials and methods

105 ### Insects

106 Adult *P. spumarius* individuals were collected by means of an entomological sweep net in July
107 2018 in wild areas in north-western Italy (Piedmont and Lombardy regions) (Table 1), where both
108 annual and perennial herbs were present, as well as tree species, all potentially hosting the
109 spittlebug. Parts from all plants where insects had been captured were collected for identification.
110 Insects were maintained at the DISAFA laboratories for at least 7-10 days in two distinct lab
111 rearings in mesh cages (580 × 580 × 600 mm) containing potted *Digitaria ciliaris* (Retz.) Koeler or

112 *Medicago sativa* L. plants, under outdoor conditions in a sheltered place. Weeds other than *D.*
113 *ciliaris* or *M. sativa* emerging from the pot soil were manually removed daily. Furthermore, other
114 adult specimens (6 males and 4 females) were picked immediately after field collection and
115 preserved at -20°C for molecular analyses (named group 5 in Figure 1).

116 **Rearing trials on different feeding substrates**

117 *P. spumarius* from the lab rearings were divided in four subgroups differing for the food source
118 (Figure 1). Insects in groups 1-3 were taken from the rearing on *D. ciliaris*, while those in group 4
119 were obtained from the rearing on *M. sativa*. Spittlebugs dedicated to group 1 were kept on *D.*
120 *ciliaris* plants and then collected for molecular analyses, whereas adults in groups 2 and 3 were
121 moved to a new feeding substrate, consisting of an artificial diet or a *M. sativa* seedling,
122 respectively, for three days. The artificial diet was prepared with 0.7 mM L-glutamine, 0.1 mM L-
123 asparagine, and 1 mM sodium citrate, pH 6.4, according to Killiny & Almeida (2009), and was
124 supplied by using an artificial feeding system as described by Gonella et al. (2015). Insect mortality,
125 integrity of the membrane containing the diet, and moisture in the feeding systems were checked
126 daily for insects maintained on the artificial diet; only live specimens were kept for molecular
127 analysis. *P. spumarius* individuals dedicated to group 4 were moved from the rearing on *M. sativa*
128 to a *D. ciliaris* seedling for 3 days. At the end of the experiments, 10 live insects were collected
129 from each group (namely, 5 males and 5 females for group 1; 6 males and 4 females for group 2; 7
130 males and 3 females from group 3; 4 males and 6 females for group 4); spittlebug whole bodies
131 were preserved at -20°C for molecular analyses.

132 **DNA extraction, PCR, and sequencing**

133 A molecular diagnostic protocol was set up to assess whether the visual host plant identification
134 could be confirmed, and if the genetic material of possible previous host plants could be retained in
135 and detectable from the gut of *P. spumarius*. Total DNA was extracted from the whole body of *P.*
136 *spumarius* samples by using the QIAamp® PowerFecal® DNA Kit (Qiagen, Italy), according to the

manufacturer's instruction, with the following modification: insect tissues were lysed and homogenized with a sterile pestle in an Eppendorf tube with 750 µl of PowerBead Solution, then the homogenate was transferred in the Bead tube to proceed with the protocol indicated by the supplier. Subsequently, the DNA was submitted to nested PCR targeting the chloroplast region between the *trnL* and *trnF* genes, by using primer pairs partially modified from Taberlet et al. (1991; 2007), optimized to cover a wider range of plants from mixed DNA sources. Specifically, direct PCR was performed with the modified forward primer Fc1 (5'-CGRAATYGGTAGACGCTACG-3') coupled with the reverse primer Rf (Taberlet et al., 1991), targeting the *trnL-trnF* region. Products of direct PCR reactions were submitted to 1:40 dilution with sterile water and amplified in nested PCR with the modified forward primer Fg1 (5'-GGGYRHTCCTGRKCCAA-3') and the reverse primer Rd (Taberlet et al. 1991), targeting the *trnL* intron. All reactions were performed 25 µl reaction mixture containing 1 × PCR buffer (Solis Biodyne, Estonia), 0.12 mM of each dNTP, 0.3 mM of each primer, 1 U of HOT FIREPol® Taq polymerase (Solis Biodyne) and 2 µl of DNA template. The cycling conditions were as follows (for both direct and nested PCRs): 95°C for 15 min; 35 cycles of 95°C for 30 s, 48°C for 45 s and 72°C for 1 min; and 7 min at 72°C. Amplicons from the nested PCRs were sequenced at Eurofin Genomics S.r.l [Via Bruno Buozzi, 2, 20090 Vimodrone (Milano), Italy]; sequences were subjected to Nucleotide BLAST analysis against nr database (https://blast.ncbi.nlm.nih.gov/Blast.cgi). For some, When direct sequencing did not allow to obtain a single clean sequence (in samples belonging to groups 2-4), amplicons were cloned using a pGEM® T-easy Vector Cloning Kit (Promega, Italy), and five clones from each amplicon were sequenced. Moreover, PCR products obtained from field-collected insects (group 5 samples) were immediately cloned as described above, and 10 clones from each amplicon were sequenced.

Results

A total of 144 *P. spumarius* adults were captured in the field both on herbs, shrubs and trees. The plants belonged to 18 species in 11 different families: Aceraceae (one species), Asteraceae (four species), Betulaceae (two species), Convolvulaceae (one species), Equisetaceae (one species), Fabaceae (three species), Fagaceae (one species), Poaceae (one species), Rosaceae: (two species), Pinaceae (one species), and Vitaceae: (one species) (Table 2).

After insect collection and rearing on different sources, we tested DNA samples by chloroplast-targeted nested PCR. After amplification of all of the 10 samples from spittlebugs fed for at least one week on *D. ciliaris* (group 1), a 450 bp-long amplicon was obtained (Figure 2); all sequences obtained from these PCR products were clean and referable to *D. ciliaris* (Table 3). These results indicated that a sufficient amount of amplifiable and clean DNA from plant chloroplast was achievable from the digestive tract of a xylem feeder such as *P. spumarius*. Furthermore, a minimum continuous exposure to a single plant of seven days was demonstrated to be sufficient to avoid the presence of "contaminant" chloroplast DNA from plants possibly consumed by the specimens before collection, allowing the experiment continuation.

As a second step we evaluated how long viable chloroplast DNA could persist in the gut of *P. spumarius*, by using an artificial feeding system containing a chloroplast-free liquid diet. High mortality was recorded in adults fed with the artificial diet: a total number of 97 specimens were individually used to obtain 10 surviving for three days, with a 89.7% mortality rate. After DNA extraction from the 10 *P. spumarius* specimens found alive after being maintained for three days on artificial diet, PCR reactions on these samples produced amplicons ranging from 200 to 300 bp. However, none of the resulting sequences were directly readable, and even after cloning only 100-150 bp-long sequences were obtained, unrelated to the target chloroplast *trnL* gene and hence referable to artefacts.

Additional trials were carried out to establish the influence of a host plant shift on the results of molecular plant identification. After PCR analysis on DNA samples from both group 3 and group 4

187 *P. spumarius* individuals (spittlebugs transferred from *D. ciliaris* to *M. sativa*, or from *M. sativa* to
188 *D. ciliaris*, respectively, Figure 1), amplicons ranging from 350 to 450 bp were always obtained
189 (Figure 2). Considering group 3 samples, 6 out of 10 sequenced amplicons were still clearly
190 referable to *D. ciliaris*, whereas the remaining four PCR products (40%) were cloned. All of the
191 Operational Taxonomic Units (OTUs) from five sequenced clones for each amplicon were related
192 to *M. sativa* (Table 3). Conversely, 8 out of 10 sequences from group 4 samples were directly
193 readable; four of them were OTUs related to *D. ciliaris* and four to *M. sativa*, while the remaining
194 two amplicons (20%) were cloned. In both cases all of the five sequenced clones were referable to a
195 single OTU; one of these sequences had as the closest relative *D. ciliaris* and the other *M. sativa*.
196 As a final result, half of the 10 samples from group 4 were related to *D. ciliaris* and half to *M.*
197 *sativa* (Table 3).

198 To assess the suitability of this protocol for the collected *P. spumarius* populations, we finally
199 tested 10 randomly selected adults directly preserved after field collection (referred to as group 5).
200 The target chloroplast gene was successfully amplified from all of the 10 samples coming from this
201 group, producing amplicons ranging from 200 to 500 bp These products were cloned and 10 clones
202 for each sample sequenced; the results of sequencing are indicated in Table 4. Overall, a total of 11
203 single plant species were identified through this method; all of them belonged to the host range
204 recorded during the field survey (Table 2). Six of the identified plant species were woody plants and
205 five were herbs. In eight samples, all of obtained OTUs were related to a single plant species, even
206 though in one case two distinct OTUs affiliated to the same species were found. In the remaining
207 two samples, OTUs related to two different species were retrieved; either belonging to tree or
208 herbaceous plants. In one sample a total of three OTUs were obtained; however, two of them were
209 referable to a single plant species. No amplicon containing OTUs affiliated to more than two plant
210 species was found.

211

Discussion

The meadow spittlebug *P. spumarius* is known to be a highly polyphagous species being able to feed on monocotyledonous and dicotyledonous plants either as nymph or adult (Ossiannilsson, 1981, Cornara et al., 2017, Cornara et al., 2018a; Di Serio et al., 2019). The identification of the proper host plants is simplified by the stationary behaviour of the nymph, as they produce and reside in a protective froth. On the contrary, the simple collection of adults on a plant species does not necessarily indicate that they had actively used that plant, due to the high mobility of the stage; for this reason, alternative methods to identify the actual host plants of single adults are required. The molecular analysis of plant-related DNA in the digestive tract of *P. spumarius* to identify its gut content confirmed previous results supporting the use of this technique on insects that feed on saps (Rodney Cooper et al., 2016). Specifically, we provided the first nested PCR-based identification of the gut content in xylem feeders. Insects with this feeding behaviour are thought to retain a very low plant-related DNA load. Hence, a major challenge is to successfully amplify potential mixed chloroplast DNA deriving from multiple host plants used by insects with such a little plant DNA concentration in their gut. Indeed, even though our newly designed primers were conceived to be universal, they may display different affinity levels with specific taxa (Bista et al., 2018; Piñol et al., 2019) resulting in biased amplification of single plant OTUs from an insect that had actually fed on many hosts. Up to three distinct OTUs were found in a single field-collected specimen, corresponding to two plant species always belonging to the observed host range, confirming that multiple hosts can be detected with this method. The relatively low number of distinct host plants detected by this method may be related to i) limited number of ingested/probed plants, or ii) differential DNA degradation based on inversely proportional time length of the feeding period on different hosts. The real number of different plant species used by an adult individual of *P. spumarius* during its life cycle is not known; however, studies on its feeding preference showed that adults may switch from an host to another in a few hours, and they may

perform several feeding events in a short time (Markheiser et al., 2020). These reports, along with the high mobility of *P. spumarius* adults, suggest that more than two plants have been used by collected adults before being sampled. On the other hand, the DNA of plants used for a limited time may have been degraded during the digestion process. Molecular analysis of chloroplast DNA from the gut of herbivorous insects feeding on seeds or roots allowed detection of the provided food source after three days of digestion (Wallinger et al., 2013; 2015); however, no indication on the real persistence of plant DNA from multiple hosts in the gut of xylem feeders is presently available. Moreover, plant identity was reported to affect post-feeding DNA detection success (Wallinger et al., 2013), suggesting that the different host plants may undergo a different fate once inside the insect gut.

We investigated the effect of a host plant switch on the newly designed molecular method. The retrieval of OTUs related to both plants sequentially provided to *P. spumarius* in experiments 3 and 4 (nearly in a 1:1 ratio) supports to the absence – or the limited presence – of PCR-related bias produced by using the proposed primer pairs, at least considering the two plant species in the families Poaceae and Fabaceae. However, we must take into account that the possible differential detection of the two hosts may be related to different amounts of ingested chloroplast DNA and/or different responses to the insect digestion, consistently with the results obtained from field-collected adults. Accordingly, in some samples the detected plant did not correspond to the last plant species being supplied, suggesting that the range of data being potentially achieved by using this protocol is not restricted to the last plant consumed by the spittlebug, but it may depend on the amount of DNA ingested by each single tested adult from the two host plants, or to the DNA quality inside the insect gut at the time of collection. The retrieval of chloroplast DNA from an earlier provided plant after host switch is consistent with previous results reported for psyllids, where the DNA of the first host plant was found up to one week after insects were moved to a different species (Rodney Cooper et al., 2016). However, in these experiments we never found OTUs related to both host plants in the

same individual sample, even when cloning was required, despite the insects had actually consecutively fed on the two species. Even though we cannot rule out the possibility that sequencing a higher number of clones from each sample would allow obtaining sequences from both the plant species, this result indicates that single PCR reactions may support the preferential amplification of an individual target sequence, and this must be taken into account for data interpretation.

Another unknown aspect, when analysing field-collected specimens, is the time lapse from the last feeding event to collection; experiment 2 was set up exposing insects to an artificial diet to establish the persistence of plant DNA in the gut of *P. spumarius*. However, very low survival rates were observed for insects used in this experiment. Previous evidences have been provided on the marked recalcitrance of *P. spumarius* to accept artificial diets as food sources (Cornara et al., 2019). It is likely that the low percentage of adults capable to survive for three days had accidentally ingested the diet after piercing the membrane by chance. Molecular analyses performed on these insects did not result in successful amplification of the chloroplast DNA of the plant used by *P. spumarius*, most probably because of partial degradation of the little amount of plant DNA, which was still present after three days in which insects were maintained on the artificial feeding systems. It must be pointed out that such a degradation was only moderately visible in the experiments involving a switch of host plant (groups 3-4); therefore our results may be either suggestive of i) a partial interference of the diet components with chloroplast DNA integrity in the insect gut, or ii) marked tissue alteration in the digestive tract of almost starving individuals. In the light of these results, we can conclude that stable chloroplast DNA can be retrieved in the gut of *P. spumarius* only after very short time from the end of feeding.

Determining the host plants of *P. spumarius* is a key step parallel to *X. fastidiosa* monitoring actions, necessary in areas where the pathogen has not been recorded yet for facing its rapid spread in Europe. The molecular tool described here can support field observations, especially in

uninfected areas, where the knowledge of the main food sources of single adults could be very useful to immediately drive the control measures in case of pathogen detection in a plant (either wild or cultivated) in that area, even before finding infected vectors. In case of detection of individuals infected by the pathogen, it will be still possible to achieve information on the plants that have been used by these infected specimens, allowing their rapid eradication. Extensive monitoring of *X. fastidiosa* infection sources is very important in north-western Italy even though the pathogen has not been recorded in this region yet, since it includes several areas classed as climatically suitable for the pathogen (EFSA, 2019). Notably, *P. spumarius* was reported as abundant (almost reaching the density of 2 adults per plant) on olive trees in north-western Italy all throughout the olive growing season, suggesting a high risk of pathogen outbreaks in case of *X. fastidiosa* invasion (Bodino et al., 2019; 2020). Moreover, in our field collections, we commonly found *P. spumarius* adults on grapevine in grapevine growing sites during summer, which might be a serious issue in case of presence of Pierce's disease, which has been recently recorded in Europe, since the spittlebug has been demonstrated to transmit the strain being the causal agent to grapevine (Moralejo et al., 2019). However, we must point out that it is impossible to cover the full host range of an individual during its entire life, because the only chloroplast DNA that can be retrieved belongs to a limited number of host plants used by *P. spumarius* in the previous few days. Hence, considering that *X. fastidiosa* is persistent in in the foregut of its vectors (Chatterjee et al., 2008), if a specimen had been infected long before collection, some plant species representing a source for pathogen acquisition may remain unidentified. For this reason, this method could be more effectively applied in early phases of adult appearance and soon after the first detection of *X. fastidiosa* infection in a specific area, when the pathogen can be presumed to have been recently acquired by putatively infected insects.

Additional application of the proposed technique may be found in investigating the ecology of *P. spumarius*. The molecular identification of the host plant(s) used by single adults can be used to

312 establish the seasonal behaviour of this polyphagous insect, which has been earlier observed to
313 move from herbaceous hosts to bushes and trees over the summer (Cornara et al., 2017). Our results
314 showed that about half of field-collected specimens contained DNA from herbaceous species, while
315 in the other half DNA from tree species was found, suggesting a transition among different host
316 types, in agreement with our field samplings and with the collection period. Indeed, adults are most
317 often found on arboreous species in mid-late summer, probably because of a typical grassland
318 reduction during the warmer season. In contrast, herbaceous host plants are most frequently
319 reported in early summer and autumn, as females lay eggs mainly on shrivelled grass near the soil,
320 since the end of August and until late season (Halkka et al., 1967; Bodino et al., 2019). However, it
321 is worth of remark that our survey was conducted in north-western Italy, which falls in the Cfa
322 climate type according to the Köppen-Geiger classification (Peel et al., 2007), displaying a
323 tempered and humid climate with hot summer. Differently, southern and coastal northern Italy,
324 where the behaviour of *P. spumarius* was previously recorded (Bodino et al., 2019; 2020), both fall
325 in the Csa type, with hot and dry summer (Peel et al., 2007). The climatic conditions of north-
326 western Italy prevent the complete desiccation of the turf, allowing the meadow spittlebug adults to
327 exploit this source also during summer, being less forced to migrate towards the woodland. A deep
328 screening of *P. spumarius* populations occurring in distinct areas may further clarify this
329 behavioural trait, to verify to what extent it is affected by differential environmental conditions
330 determining plant species composition, and consequently food and shelter availability.

331 In conclusion, in this work we provided new insights on the identification of host plants of *P.*
332 *spumarius* adults, contributing to elucidate its feeding behaviour in north-western Italy, where it
333 may become a serious threat in case of introduction of *X. fastidiosa*, as this region includes many
334 olive- and grapevine-growing areas. Furthermore, we demonstrated for the first time that the
335 feeding behaviour of a xylem feeding insect such as *P. spumarius* sustains the amplification of plant

DNA. Our results provide a useful tool for better understanding the spread of *X. fastidiosa* in just invaded areas, contributing to the study of vector ecology and disease epidemiology.

Acknowledgements

The authors are grateful to Enrico Busato, Marco Pansa and Carolina Pisa for helping in insect collection. The study was supported by the PRECONFITOLOMB project (Prevenzione e controllo di nuove emergenze fitosanitarie in Lombardia*: Popillia japonica* e *Xylella fastidiosa*), funded by Regione Lombardia.

References

Almeida RPP, Blua MJ, Lopes JRS & Purcell AH (2005) Vector transmission of *Xylella fastidiosa*: applying fundamental knowledge to generate disease management strategies. Annals of the Entomological Society of America 98: 775-86.

Avanesyan A (2014) Plant DNA detection from grasshopper guts: a step-by-step protocol, from tissue preparation to obtaining plant DNA sequences. Applications in Plant Sciences 2: 1300082.

Baldi P & La Porta N (2017) *Xylella fastidiosa*: host range and advance in molecular identification techniques. Frontiers in Plant Science 8: 944.

Bista I, Carvalho GR, Tang M, Walsh K, Zhou X et al. (2018) Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. Molecular Ecology Resources 18: 1020-1034.

357 Bodino N, Cavalieri V, Dongiovanni C, Plazio E, Saladini MA et al. (2019) Phenology, seasonal
358 abundance and stage-structure of spittlebug (Hemiptera: Aphrophoridae) populations in olive
359 groves in Italy. Scientific Reports 9: 17725.

360 Bodino N , Cavalieri V, Dongiovanni C, Saladini MA, Simonetto A et al. (2020) Spittlebugs of
361 Mediterranean olive groves: host-plant exploitation throughout the year. Insects 11: 130.

362 Cavalieri V, Altamura G, Fumarola G, di Carolo M, Saponari M et al. (2019) Transmission of
363 *Xylella fastidiosa* subspecies *pauca* sequence type 53 by different insect species. Insects 10:
364 324.

365 Chatterjee S, Almeida RPP & Lindow S (2008) Living in two worlds: the plant and insect lifestyles
366 of *Xylella fastidiosa*. Annual Review of Phytopathology 46: 243-271.

367 Cordeiro AB, Sugahara VH, Stein B & Leite Junior RP (2014) Evaluation by PCR of *Xylella*
368 *fastidiosa* subsp. pauca transmission through citrus seeds with special emphasis on lemons
369 (*Citrus limon* (L.) Burm. f). Crop Protection 62: 86-92.

370 Cornara D, Sicard A, Zeilinger AR, Porcelli F, Purcell AH et al. (2016) Transmission of *Xylella*
371 *fastidiosa* to grapevine by the meadow spittlebug. Phytopathology 106: 1285-1290.

372 Cornara D, Saponari M, Zeilinger AR, de Stradis A., Boscia, D et al. (2017) Spittlebugs as vectors
373 of *Xylella fastidiosa* in olive orchards in Italy. Journal of Pest Science 90: 521-530.

374 Cornara D, Bosco D & Fereres A (2018a). *Philaenus spumarius*: when an old acquaintance
375 becomes a new threat to European agriculture. Journal of Pest Science 91: 957.

376 Cornara D, Garzo E, Morente M, Moreno A, Alba-Tercedor J et al. (2018b) EPG combined with
377 micro-CT and video recording reveals new insights on the feeding behavior of *Philaenus*
378 *spumarius*. PLoS ONE 13 : e0199154.

379 Cornara D, Ripamonti M, Morente M, Garzo E, Bosco D, et al. (2019) Artificial diet delivery
380 system for *Philaenus spumarius*, the European vector of *Xylella fastidiosa.* Journal of Applied
381 Entomology: 143: 882–892.

382 De la Cadena G, Papadopoulou A, Maes J. & Gómez-Zurita J (2017) Evaluation of bias on the
383 assessment of diet breadth of herbivorous insects using molecular methods. Insect Science 24:
384 194-209.

385 Di Serio F, Bodino N, Cavalieri V, Demichelis S, Di Carolo M et al. (2019) Collection of data and
386 information on biology and control of vectors of *Xylella fastidiosa*. EFSA supporting
387 publication 2019:EN-1628.

388 Dongiovanni C, Cavalieri V, Bodino N, Tauro D, Di Carolo M et al (2019) Plant selection and
389 population trend of spittlebug immatures (Hemiptera: Aphrophoridae) in olive groves of the
390 Apulia region of Italy. Journal of Economic Entomology 112: 67–74.

391 European Food Safety Authority (EFSA) (2018) Update of the *Xylella* spp. host plant database.
392 EFSA Journal 16: 5408.

393 EFSA Panel on Plant Health (PLH), Bragard C, Dehnen-Schmutz K, Di Serio F, Gonthier P,
394 Jacques M-A et al (2019) Update of the scientific opinion on the risks to plant health posed by
395 *Xylella fastidiosa* in the EU territory. EFSA Journal 17: 5665.

396 Gariepy TD, Kuhlmann U, Gillott C & Erlandson M (2007) Parasitoids, predators and PCR: the use
397 of diagnostic molecular markers in biological control of Arthropods. Journal of Applied
398 Entomology 131: 225–240.

399 Gonella E, Pajoro M, Marzorati M, Crotti E, Mandrioli M et al (2015) Plant-mediated interspecific
400 horizontal transmission of an intracellular symbiont in insects. Scientific Reports 5: 15811.

401 Halkka O, Raatikainen M, Vasarainen A & Heinonen L (1967) Ecology and ecological genetics of
402 *Philaenus spumarius* (L.) (Homoptera). Annales Zoologici Fennici 4: 1–18.

403 Jurado-Rivera JA, Vogler AP, Reid CAM, Petitpierre E & Gómez-Zurita J (2009) DNA barcoding
404 insect–host plant associations. Proceedings of the Royal Society B: Biological Sciences 276:
405 639–648.

406 Killiny N& Almeida RPP (2009) Host structural carbohydrate induces vector transmission of a
407 bacterial plant pathogen. Proceedings of the National Academy of Sciences of the United
408 States of America 106: 22416–22420.

409 Krugner R, Sisterson MS, Backus EA, Burbank LP & Redak RA (2019) Sharpshooters: a review of
410 what moves *Xylella fastidiosa*. Austral Entomology 58: 248– 267.

411 Marcelletti S & Scortichini M (2016) Genome-wide comparison and taxonomic relatedness of
412 multiple *Xylella fastidiosa* strains reveal the occurrence of three subspecies and a new *Xylella*
413 species. Archives of Microbiology 198: 803-812.

414 Markheiser A, Cornara D, Fereres A & Maixner M (2020) Analysis of vector behavior as a tool to
415 predict *Xylella fastidiosa* patterns of spread. Entomologia Generalis 40: 1-13.

416 Matheson CD, Muller GC, Junnila A, Vernon K, Hausmann A et al. (2008) A PCR method for
417 detection of plant meals from the guts of insects. Organisms Diversity & Evolution 7: 294-303.

418 Miranda MP, Fereres A, Appezzato-da-Gloria B & Lopes JRS (2009) Characterization of electrical
419 penetration graphs of *Bucephalogonia xanthophis*, a vector of *Xylella fastidiosa* in citrus.
420 Entomologia Experimentalis et Applicata 130: 35-46.

421 Moralejo E, Borràs D, Gomila M, Montesinos M, Adrover F et al. (2019) Insights into the
422 epidemiology of Pierce's disease in vineyards of Mallorca, Spain. Plant Pathology 68: 1458–
423 1471.

Ossiannilsson F (1981) The Auchenorrhyncha (Homoptera) of Fennoscandia and Denmark. Part 2: the families Cicadidae, Cercopidae, Membracidae, and Cicadellidae (excl. Deltocephalinae). Fauna Entomologica Scandinavica 7: 223–593.

Pearson CC, Backus EA, Shugart HJ & Munyaneza JE (2014) Characterization and correlation of EPG waveforms of *Bactericera cockerelli* (Hemiptera: Triozidae): variability in waveform appearance in relation to applied signal. Annals of the Entomological Society of America 107: 650–66.

Peel MC, Finlayson BL, & McMahon TA (2007) Updated world map of the Köppen-Geiger climate classification. Hydrology and Earth System Sciences, 11: 1633–1644.

Piñol J, Senar MA & Symondson WOC (2019) The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. Molecular Ecology 28: 407–419.

Pompanon F, Daegle BE, Symondson WOC, Brown DS, Jarman SN et al. (2012) Who is eating what: diet assessment using next generation sequencing. Molecular Ecology 21: 1931-1950.

Pumariño L, Alomar O & Agustí N (2011) Development of specific ITS markers for plant DNA identification within herbivorous insects. Bulletin of Entomological Research 101: 271-276.

Redak RA, Purcell AH, Lopes JRS, Blua MJ, Mizell III RF et al. (2004) The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. Annual Review of Entomology 49: 243-270.

Rodney Cooper W, Horton DR, Unruh TR & Garczynski SF (2016) Gut content analysis of a phloem-feeding insect, *Bactericera cockerelli* (Hemiptera: Triozidae). Environmental Entomology 45: 938-44.

446 Santoiemma G, Tamburini G, Sanna F, Mori N & Marini L (2019) Landscape composition predicts
447 the distribution of *Philaenus spumarius*, vector of *Xylella fastidiosa*, in olive groves. Journal of
448 Pest Science 92: 1101–1109.

449 Sheppard SK & Harwood JD (2005) Advances in molecular ecology: tracking trophic links through
450 predator–prey food-webs. Functional Ecology 19: 751–762.

451 Sicard A, Zeilinger AR. Vanhove M, Schartel TE, Beal DJ et al. (2018) *Xylella fastidiosa*: insights
452 into an emerging plant pathogen. Annual Review of Phytopathology 56: 181-202.

453 Strona G, Carstens CJ & Beck PSA (2017) Network analysis reveals why *Xylella fastidiosa* will
454 persist in Europe. Scientific Reports 7: 71.

455 Taberlet P, Gielly L, Pautou G & Bouvet J (1991) Universal primers for amplification of three non-
456 coding regions of chloroplast DNA. Plant Molecular Biology 17: 1105-1109.

457 Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C et al. (2007) Power and limitations of the
458 chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Research 35: e14.

459 Wallinger C, Staudacher K, Schallhart N, Peter E, Dresch P et al. (2013) The effect of plant identity
460 and the level of plant decay on molecular gut content analysis in a herbivorous soil insect.
461 Molecular Ecology Resources 13: 75–83.

462 Wallinger C, Sint D, Baier F, Schmid C, Mayerand R et al. (2015) Detection of seed DNA in
463 regurgitates of granivorous carabid beetles. Bulletin of Entomological Research 105: 728-735.

464

465 **Table 1** Sampling sites of field collected adult *P. spumarius* used in this study.

Municipality	**Region, Province**	**Gps coordinates WGS84 (EPSG:4326)**
Casellette	Piedmont, Turin	45°05'50.3"N 7°28'14.6"E
Chieri	Piedmont, Turin	45°01'12.0"N 7°47'13.1"E
Grugliasco	Piedmont, Turin	45°04'01.4"N 7°35'29.4"E
Lonate Pozzolo	Lombardy, Milan	45°35'28.2"N 8°42'40.0"E
Morimondo	Lombardy, Milan	45°20'57.2"N 8°55'37.1"E

466

467

468 **Table 2** Number of adult *P. spumarius* collected in this study on different host plants.

Plant species	No. collected specimens
Herbaceous host plant	
Achillea millefolium L.	7
Arrhenatherum elatius (L.) P. Beauv. ex J. Presl & C.	4
Cirsium arvense (L.) Scopoli	11
Convolvolus arvensis L.	2
Equisetum arvense L.	1
Medicago sativa L.	15
Rubus ulmifolius Schott	1
Solidago gigantea Aiton	3
Taraxacum officinale Weber ex F.H. Wigg	10
Trifolium pratense L.	5
Total herbaceous hosts	**59**
Woody host plant	
Acer campestre L.	17
Carpinus betulus L.	5
Corylus avellana L.	9
Picea pungens Engelmann	18
Prunus avium L.	16
Quercus rubra L.	3
Robinia pseudoacacia L.	1
Vitis vinifera L.	16
Total woody hosts	**85**
Total number	**144**

469

470

471 **Table 3** Results of sequencing of chloroplast *trnL* gene amplicons from *P. spumarius* adults reared
472 on different food substrates, according to insect groups described in Figure 1. The number of plant-
473 related OTUs is indicated for each insect sample, while for each OTU the ratio between the number
474 of sequences and the total number of tested insect for each group is reported.

475

Food source (Insect group)	**No. retrieved plant OTUs**	**OTU proportion**	**Closest relative (NCBI Accession Number)**
D. ciliaris (1)	1	10/10	*Digitaria ciliaris* (LC118761)
D. ciliaris to artificial diet (2)	0	0/12	-
D. ciliaris to *M. sativa* (3)	2	6/10	*Digitaria ciliaris* (LC118761)
		4/10	*Medicago sativa* (KP174818)
M. sativa to *D. ciliaris* (4)	2	5/10	*Digitaria ciliaris* (LC118761)
		5/10	*Medicago sativa* (KP174818)

476

477

Table 4 Results of sequence analysis obtained after cloning the chloroplast *trnL* gene amplified from field-collected *P. spumarius* adults (indicated as insect group 5 in Figure 1). The number of plant-related OTUs is indicated for each insect sample, while for each OTU the ratio between the number of sequences and the total sequenced clones for each insect sample (OTU proportion) is reported as well as the obtained sequence length.

Insect sample ID	No. plant OTUs	OTU proportion	Sequence length (bp)	Closest relative (NCBI Accession Number)	% sequence identity
Ps18.1	1	10/10	310	*Corylus avellana* (KF718348)	100% (310/310 bp)
Ps18.2	1	10/10	246	*Convolvulus arvensis* (MF621879)	100% (246/246 bp)
Ps18.3	2	8/10	472	*Picea pungens* (EF440560)	99% (467/472 bp)
		2/10	471	*Plantago lanceolata* (AY101952)	100% (471/471 bp)
Ps18.4	2	7/10	396	*Achillea millefolium* (EU128988)	99%(395/396 bp)
		3/10	396	*Achillea millefolium* (EU128988)	99% (391/396 bp)
Ps18.5	1	10/10	423	*Acer campestre* (KU522504)	100% (423/423 bp)
Ps18.6	1	10/10	303	*Equisetum arvense* (GQ428069)	99% (301/303 bp)
Ps18.7	1	10/10	327	*Carpinus betulus* (AF327579)	99% (326/327 bp)
Ps18.8	3	4/10	312	*Arrhenatherum elatius* (MH569076)	100% (312/312 bp)
		2/10	312	*Arrhenatherum elatius* (MH569076)	99% (308/312 bp)
		4/10	312	*Convolvulus arvensis* (KC786130)	100% (312/312 bp)
Ps18.9	1	10/10	513	*Robinia pseudoacacia* (NC_026684)	99% (510/513 bp)
Ps18.10	1	10/10	223	*Quercus rubra* (KU186951)	99% (222/223 bp)

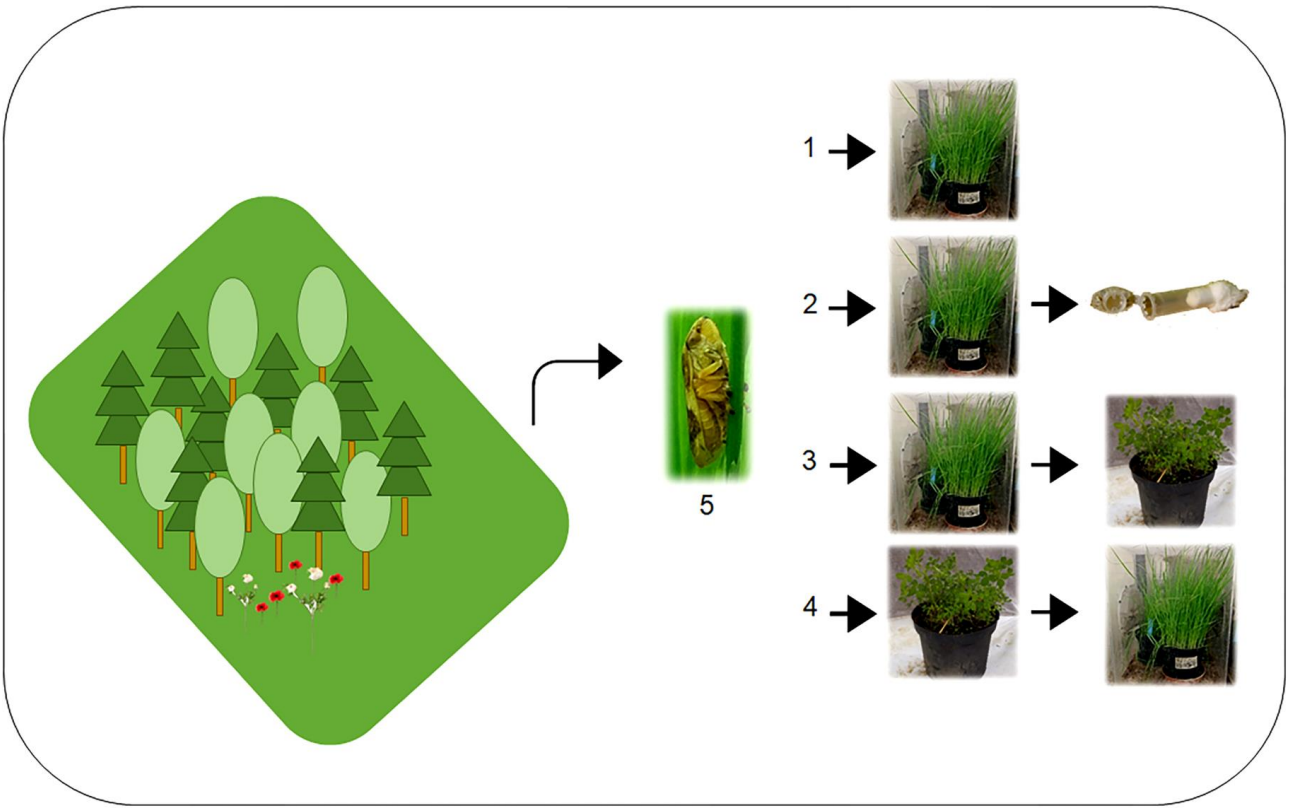

Figure 1 Experimental plan used for this work. Adult *P. spumarius* individuals were collected from meadows close to forest areas in north-western Italy and then reared on potted plants of *D. ciliaris* (1-3) or *M. sativa* (4) for at least 7 days. Afterwards, a first group of spittlebugs from the *D. ciliaris* rearing was directly collected and submitted to molecular analysis (1); whereas a second group of individuals was transferred on an artificial feeding system for 3 days (2). Additionally, a further group of specimens from *D. ciliaris* was moved to potted *M. sativa* plants (3), while the spittlebugs from the *M. sativa* rearing were transferred to *D. ciliaris* isolated plants (4) for other 3 days. A final group of insects was preserved for molecular analysis directly after field collection for the validation of the diagnostic method and for final persistence assessment of chloroplast DNA in the gut of *P. spumarius* (5).

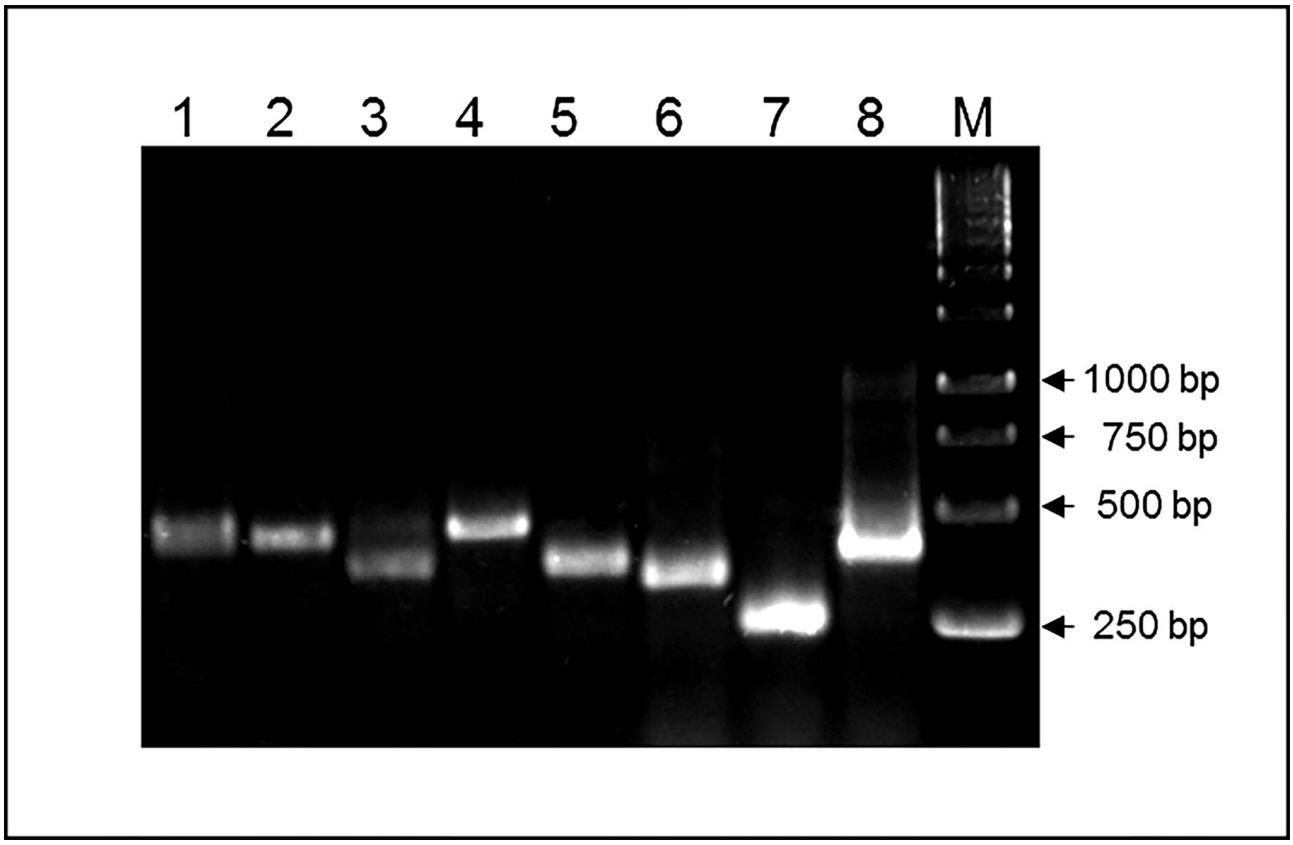

Figure 2 Electrophoresis of chloroplast *trnL* gene amplicons obtained from one *P. spumarius* adult reared for at least 7 days on *D. ciliaris* (1); two specimens maintained on *D. ciliaris* and then moved to *M. sativa* (2-3); two specimens maintained on *M. sativa* and then moved to *D. ciliaris* (4-5); and three field collected individuals (6-8), shown in Table 1 as Ps18.1 - Ps18.3. Sequences obtained from amplicons 1, 2, and 4 were referable to *D. ciliaris* (450 bp); sequences from amplicons 3 and 5 were referable to *M. sativa* (350 bp); whereas sequences from field-collected specimens (amplicons 6-8) are indicated in Table 1.